CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 620

A Local Variation Method for Bilevel Nash Equilibrium Problems

Francesco Caruso, Maria Carmela Ceparano, and Jacqueline Morgan

July 2021

University of Naples Federico II | University of Salerno | Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

Working Paper no. 620

A Local Variation Method for Bilevel Nash Equilibrium Problems

Francesco Caruso*, **Maria Carmela Ceparano**†, **and Jacqueline Morgan** ‡

Abstract

We address the numerical approximation of bilevel problems consisting of one Nash equilibrium problem in the upper level and another Nash equilibrium problem in the lower level. These problems, widely employed in engineering and economic applications, are a generalization of the well-known Stackelberg (or bilevel optimization) problem. In this paper, we define a numerical method for bilevel Nash equilibrium problems where in the lower level there is a ratio-bounded game (introduced in Caruso, Ceparano, Morgan [*CSEF Working Papers*, 593 (2020)]) and in the upper level there is a potential game (introduced in Monderer, Shapley [*Games Econ. Behav.*, 14 (1996)]). The method, relying on a derivative-free unconstrained optimization technique called local variation method, is shown to globally converge towards a solution of the problem and also allows to obtain error estimations.

Keywords: Bilevel Nash equilibrium problem, Stackelberg problem, multi-leader-follower game, ratio-bounded game, potential game, Nash equilibrium, existence and uniqueness, local variation method, global convergence, error estimation.

* Università di Napoli Federico II. E-mail: francesco.caruso@unina.it

† Università di Napoli Federico II. E-mail: mariacarmela.ceparano@unina.it

‡ Università di Napoli Federico II and CSEF. E-mail: morgan@unina.it

Table of contents

1 Introduction

Bilevel problems involving a Nash equilibrium problem both in the upper level and in the lower level are broadly used to model a wide range of situations in engineering and economics frameworks; for example in electric power markets ([9, 26, 29, 24, 2]), networks ([52, 30]), forward markets ([1, 48]), two-period Cournot competitions ([44, 41]), contract theory ([40, 14]). This kind of hierarchical structure, which dates back to the multi-leader-follower model proposed in [46], represents a natural generalization of the bilevel (or two-level) optimization problem where an optimization problem appears in the upper level and another one with a unique solution appears in the lower level (see, for example, [8, 4, 33]). Such a problem is also referred as Stackelberg problem (see, for example, [47, 37, 3]), since originally introduced by H. von Stackelberg in 1934 ([51]) to analyze single leader-follower duopolies. Concerning the issue of numerical approximation, whilst a large amount of results have been obtained for the latter problem (see, for example, [13] for approximation and regularization methods, [20, Chapter 6] and [21, Section 20.6] for solution algorithms), literature on numerical methods for bilevel problems involving equilibrium problems in both levels is more recent and less developed due to the higher complicated nature of the problem. As far as we know: in [49] a sequential nonlinear complementarity method is proposed for finding stationary points of equilibrium problems with equilibrium constraints; in [19] a sample average approximation method is shown to converge with probability 1 in stochastic multi-leader-follower oligopolies; in [28] a forward-backward splitting method has been applied for a particular class of multi-leader-follower games; in [25] the unique solution to quadratic multi-leader-follower games is approached via a smoothing technique combined with gradient and Newton methods.

In this paper, our aim is to define a manageable numerical method which guarantees global convergence and error estimations for bilevel Nash equilibrium problems. In order to achieve such a goal, we will deal with a class of bilevel problems where the lower level problem is non-parametric. More precisely, we consider

- a two-player game $\Omega \coloneqq \{J, Y_1, Y_2, f_1, f_2\}$ in the lower level, where $J = \{1,2\}$, Y_j is the strategy set of player j and $f_j \colon Y_1 \times Y_2 \to \mathbb{R}$ is the payoff function of player j for any $j \in J$;
- a parametric two-player game $\Theta_y \coloneqq \{I, X_1, X_2, l_1(\cdot, y), l_2(\cdot, y)\}$ in the upper level, where $y \in Y_1 \times Y_2$, $I = \{1,2\}$, X_i is the strategy set of player i and $l_i \colon X_1 \times X_2 \times Y_1 \times Y_2 \to \mathbb{R}$ is the payoff function of player i for any $i \in I$.

Called $X = X_1 \times X_2$ and $Y = Y_1 \times Y_2$, we are interested in the following *bilevel Nash equilibrium problem*

$$(\mathrm{BNP}) \begin{cases} \text{Find } (\bar{x}, \bar{y}) \in X \times Y \\ \text{such that } \bar{x} \text{ is a Nash equilibrium of } \Theta_{\bar{y}} \\ \text{where } \bar{y} \text{ is a Nash equilibrium of } \Omega \end{cases}$$

If $(\bar{x}, \bar{y})$ is a solution to (BNP) then it satisfies, by definition of Nash equilibrium,

$$\bar{x}_i \in \operatorname*{Arg\,min}_{x_i \in X_i} l_i(x_i, \bar{x}_{-i}, \bar{y}) \quad \text{for any } i \in I,$$

$$\bar{y}_j \in \operatorname*{Arg\,min}_{y_j \in Y_j} f_j(y_j, \bar{y}_{-j}) \quad \text{for any } j \in J,$$

where $\{-i\} = I \setminus \{i\}$ and $\{-j\} = J \setminus \{j\}$ (for example, both (y_1, y_{-1}) and (y_2, y_{-2}) mean (y_1, y_2)). The solution concept above presented naturally arises in a game theory perspective. In fact, a solution to (BNP) generates a subgame perfect Nash equilibrium (introduced in [45], see also for example [34]) of a two-stage game where the two upper-level players act as leaders and the two lower-level players react as followers. More precisely, if $(\bar{x}, \bar{y})$ is a solution to (BNP), then the strategy profile $(\bar{x}, \bar{\varphi}(\cdot))$ where $\bar{\varphi}(x) = \bar{y}$ for any $x \in X$ is a subgame perfect Nash equilibrium of such a two-stage game. Moreover, we emphasize that the concept of solution to (BNP) extends the Stackelberg strategy pair concept defined in [31, Definition 1.1] (also denoted as Stackelberg equilibrium in [32]) for Stackelberg problems. Hence, from now on, the upper-level players and the lower-level players in (BNP) are referred as the leaders and the followers, respectively.

Remark 1.1 As regards the comparisons with other existing solution concepts, we point out that the set of solutions to (BNP) include the set of solutions to equilibrium problems with equilibrium constraints (EPEC for short, see [49, 50] for definitions), meaning that, when adapting the solution concept in [49] to our framework, it represents a particular selection of our concept (note that the existence of EPECs' solutions is guaranteed only for specific classes of problems, as for example in [44, 1, 48]). Instead, in general the solution to (BNP) is not connected with the L/F Nash equilibrium introduced in [42] for multi-leader-follower games.

When the lower-level game has a unique Nash equilibrium, the theoretical procedure to determine a solution to (BNP) is straightforward: first one finds the Nash equilibrium of Ω, then one replaces it in the upper-level game and solves the resulting Nash equilibrium problem. Nevertheless, the numerical procedure to determine a solution to (BNP) remains a crucial issue since the analytic expression of the lower-level Nash equilibrium is generally not available.

In this paper, we introduce a numerical method, called *bilevel local variation method* (BLVM), that globally converges to a solution to (BNP) for a class of bilevel Nash equilibrium problems where the uniqueness of the lower-level Nash equilibrium is ensured. More precisely, we consider bilevel Nash equilibrium problems such that the lower-level game is ratio-bounded (see [11] where the class of ratio-bounded games has been introduced in order to investigate the global convergence towards Nash equilibria of the affine relaxations of the best response algorithm). Such a ratio-boundedness property actually guarantees that the lower-level game has a unique Nash equilibrium. Moreover, the upper-level game Θ_y is assumed to be a potential game (see [35]) for any $y \in Y$, as frequently considered in many applications and computations. The BLVM is defined by making use of a derivative-free optimization technique, called *local variation method* (LVM), introduced in [15] for variational problems and exploited in [16] for function minimization problems, in [17, 36, 18] for zero-sum games and in [10] for non-zero-sum games. In particular such a technique allows us to obtain also error estimations for BLVM.

The paper is structured as follows. Section 2 concerns the issue of the existence and uniqueness of a solution to (BNP): we state the assumptions on the games in the lower and upper level, recall the class of ratio-bounded games and show relevant properties of problem (BNP). Section 3 provides the key preparatory results for the approximation of the solution to (BNP). Section 4 is devoted to the main purpose of the paper: the numerical approximation of the solution to (BNP). First, we recall the LVM and its associated convergence analysis (Section 4.1); then we define the BLVM and we show its global convergence towards the solution to (BNP) together with error estimations and rate of convergence (Section 4.2).

2 Existence and uniqueness of a solution to (BNP)

In this section we illustrate the assumptions on the games in the upper level and in the lower level which allow to show the existence and the uniqueness of a solution to (BNP).
Throughout the paper we assume that $X_i = \mathbb{R}^{p_i}$ for any $i \in I$ and $Y_j = \mathbb{R}^{q_j}$ for any $j \in J$, so $X = \mathbb{R}^p$ and $Y = \mathbb{R}^q$ with $p = p_1 + p_2$ and $q = q_1 + q_2$, and denote with $\|\cdot\|$ the Euclidean norm and with $\langle \cdot, \cdot \rangle$ the usual inner product.

Let us start with the lower-level game $\Omega = \{J, Y_1, Y_2, f_1, f_2\}$.

($\mathcal{F}_1$) f_j is strongly convex on Y_j uniformly on Y_{-j}, for any $j \in J$;
i.e. there exists $m_j > 0$ such that for any $y_j', y_j'' \in Y_j$, any $y_{-j} \in Y_{-j}$ and any $s \in [0, 1]$

$$f_j(sy_j' + (1-s)y_j'', y_{-j}) \leq sf_j(y_j', y_{-j}) + (1-s)f_j(y_j'', y_{-j}) - m_j s(1-s)\|y_j' - y_j''\|^2, \text{ for any } j \in J.$$

By assumption ($\mathcal{F}_1$), the function $b_j \colon Y_{-j} \to Y_j$

$$\{b_j(y_{-j})\} = \operatorname*{Arg\,min}_{y_j \in Y_j} f_j(y_j, y_{-j}) \tag{1}$$

is well-defined for any $j \in J$ and the Nash equilibria of Ω can be characterized in terms of fixed points of the function $b \colon Y_1 \to Y_1$ defined by

$$b \coloneqq b_1 \circ b_2, \tag{2}$$

that is $(\bar{y}_1, \bar{y}_2) \in Y$ is a Nash equilibrium of Ω if and only if $\bar{y}_1 = b(\bar{y}_1)$ and $\bar{y}_2 = b_2(\bar{y}_1)$.

Remark 2.1 Assumption ($\mathcal{F}_1$) is properly more demanding than the strong convexity of $f_j(\cdot, y_{-j})$ and it is not connected with the convexity of f_j on Y (see [10, Remark 4.1] for further discussion). Note that the well-definedness of b_j in (1), as well as all the results that will be shown in this and the next section, hold even by assuming just the strong convexity of $f_j(\cdot, y_{-j})$. We chose to state at once the assumption ($\mathcal{F}_1$) for the sake of readability, since it will play a role in the numerical approximation analysis.

($\mathcal{F}_2$) f_j is twice continuously differentiable on Y, the Hessian matrix $D^2_{y_j} f_j(y) \in \mathbb{R}^{p_j \times p_j}$ is invertible for any $y \in Y$ and

$$\lambda_j \coloneqq \sup_{y \in Y} \|[D^2_{y_j} f_j(y)]^{-1} \cdot D_{y_{-j}}(D_{y_j} f_j)(y)\| \in \mathbb{R}, \tag{3}$$

for any $j \in J$.

Assumption ($\mathcal{F}_2$) allows to define, for any $y_1', y_1'' \in Y_1$ and $y_2 \in Y_2$, the $q_1 \times q_1$ matrix

$$H(y_1', y_1'', y_2) = [D^2_{y_1} f_1(y_1', y_2)]^{-1} \cdot D_{y_2}(D_{y_1} f_1)(y_1', y_2) \cdot [D^2_{y_2} f_2(y_1'', y_2)]^{-1} \cdot D_{y_1}(D_{y_2} f_2)(y_1'', y_2), \tag{4}$$

which is needful for the definition of the class of games we are interested in the lower level.

Definition 2.1 [11, Definition 2.1] Let $(\alpha, \beta) \in \mathbb{R}^2$. The game Ω is (α, β)-ratio-bounded if for any $y_1', y_1'' \in Y_1$ and $y_2 \in Y_2$

$$\alpha \leq \frac{y_1^T[H(y_1', y_1'', y_2)]y_1}{\|y_1\|^2} \leq \beta \quad \text{for any } y_1 \in Y_1 \setminus \{0\}.$$

The class of (α, β)-ratio-bounded games contains games widely used in literature as for example quadratic potential games, quadratic zero-sum games and non-quadratic games with a bilinear strategic term (see [11, Section 3] for further discussion). For such a class, the existence of a unique Nash equilibrium has been proved, suitable affine relaxations of the best response dynamics have been shown to globally converge towards the Nash equilibrium and related error estimations have been obtained; this will be argued more in detail in the next section.
Moreover, assumption $(\mathcal{F}_2)$ guaranteed the well-definedness of the real number

$$\lambda \coloneqq \lambda_1 \lambda_2, \tag{5}$$

and of the set

$$S_\lambda \coloneqq \{(\alpha, \beta) \in \mathbb{R}^2 \mid \alpha > 1 \text{ or } \min\{\beta, \lambda\} < 1\}. \tag{6}$$

which are crucial for the statement of the following assumption and for the uniqueness of the Nash equilibrium of Ω.

$(\mathcal{G}_1)$ Ω is (α, β)-ratio-bounded with $(\alpha, \beta) \in S_\lambda$.

Proposition 2.1. *Assume $(\mathcal{F}_1)$-$(\mathcal{F}_2)$ and $(\mathcal{G}_1)$. Then Ω has a unique Nash equilibrium $(\bar{y}_1, \bar{y}_2)$, where $\bar{y}_1$ is the unique fixed point of $b_1 \circ b_2$ and $\bar{y}_2 = b_2(\bar{y}_1)$.*

Proof. The thesis follows by applying results showed by the authors in [10, 11], which exploit contraction properties of suitable affine combinations involving the function b. More precisely, depending on the value of $\lambda \in \mathbb{R}$, the set S_λ in (6) becomes

$$S_\lambda = \begin{cases} \mathbb{R}^2, & \text{if } \lambda < 1 \\ \{(\alpha, \beta) \in \mathbb{R}^2 \mid \alpha > 1\} \cup \{(\alpha, \beta) \in \mathbb{R}^2 \mid \beta < 1\}, & \text{if } \lambda \geq 1. \end{cases}$$

Hence, we distinguish two cases: $\lambda < 1$ and $\lambda \geq 1$.
When $\lambda < 1$, the result holds by [11, Theorem 4.1(i)]. When $\lambda \geq 1$, since it must be $\alpha \leq \beta$ by Definition 2.1, only one of the following situations happens: Ω is (α, β)-ratio-bounded with $\alpha > 1$ or Ω is (α, β)-ratio-bounded with $\beta < 1$. In the first situation, the thesis follows from [10, remarks 2.1 and 2.8 and Theorem 2.10]. In the second one, from [11, Theorem 4.4(i)]. $\square$

Remark 2.2 The set S_λ in (6) plays a key-role for the uniqueness result above illustrated. In fact, if Ω is (α, β)-ratio-bounded with $(\alpha, \beta) \in \mathbb{R}^2 \setminus S_\lambda$ (i.e. $\alpha \leq 1 \leq \min\{\beta, \lambda\}$), then in general neither the existence nor the uniqueness of Nash equilibria of Ω can be guaranteed; see [11, examples 4.4 to 4.6].

Let us consider now the parametric upper-level game $\Theta_y = \{I, X_1, X_2, l_1(\cdot, y), l_2(\cdot, y)\}$, where $y \in Y$.

$(\mathcal{G}_2)$ for any $y \in Y$, the game Θ_y is a potential game with potential $P(\cdot, y) \colon X \to \mathbb{R}$.

By definition of potential game (see [35]), assumption ($\mathcal{G}_2$) means that for any $y \in Y$

$$l_i(x_i', x_{-i}, y) - l_i(x_i'', x_{-i}, y) = P(x_i', x_{-i}, y) - P(x_i'', x_{-i}, y)$$

for any $x_i', x_i'' \in X_i$, $x_{-i} \in X_{-i}$ and $i \in I$. For characterizations of the potential games see [22, 7], whereas for the relations between Nash equilibria of potential games and minimizers of the potential see [35, 38, 12].
The next assumption involves directly the potential P.

($\mathcal{L}_1$) P is strongly convex on X uniformly on Y;
i.e. there exists $m_P > 0$ such that for any $x', x'' \in X$, any $y \in Y$ and any $s \in [0,1]$

$$P(sx' + (1-s)x'', y) \leq sP(x', y) + (1-s)P(x'', y) - m_P\, s(1-s)\|x' - x''\|^2.$$

By assumption ($\mathcal{L}_1$), the function $r\colon Y \to X$

$$\{r(y)\} = \operatorname*{Arg\,min}_{x \in X} P(x, y), \tag{7}$$

is well-defined and the Nash equilibria of Θ_y are connected with $r(y)$, as provided in the following result.

Proposition 2.2. *Assume ($\mathcal{G}_2$), ($\mathcal{L}_1$) and that $P(\cdot, y)$ is continuously differentiable on X for any $y \in Y$. Then, for any $y \in Y$ the game Θ_y has a unique Nash equilibrium, which is $r(y)$.*

Proof. Let $y \in Y$. It is worth to preliminarily note that the set of Nash equilibria of Θ_y coincides with the set of equilibria of $\{I, X_1, X_2, P(\cdot, y), P(\cdot, y)\}$ and that any miminizer of $P(\cdot, y)$ is a Nash equilibrium of Θ_y, by [35, Lemma 2.1]. By the first order characterization of the convexity of $P(\cdot, x_2, y)$ and $P(x_1, \cdot, y)$, the set of Nash equilibria of $\{I, X_1, X_2, P(\cdot, y), P(\cdot, y)\}$ equals the set $Z \coloneqq \{x \in X \mid (D_{x_1}P(x, y), D_{x_2}P(x, y)) = (0, 0)\}$. Moreover, assumption ($\mathcal{L}_1$) and the differentiability of $P(\cdot, y)$ guarantee that $Z = \{r(y)\}$, as $r(y)$ is the unique minimizer of $P(\cdot, y)$ over X. Therefore $r(y)$ is the unique Nash equilibrium of the game Θ_y. □

Remark 2.3 We point out that the result on the uniqueness of the Nash equilibrium of Θ_y in Proposition 2.2 cannot be obtained either via [43, Theorem 2], since X_1 and X_2 are not compact, or via [38, Corollary p.226], as $l_1(\cdot, y)$ and $l_2(\cdot, y)$ are not assumed to be bounded. Furthermore, the continuous differentiability of $P(\cdot, y)$ in Proposition 2.2 cannot be dropped, as shown in the following example based on [38, Remarks].

Example 2.1 Let $X = \mathbb{R}^2$ and $P(x, y) = x_1^2 + x_2^2 + 2|x_2 - x_1|$. The function $P(\cdot, y)$ has a unique minimizer over $\mathbb{R}^2$, namely $(0, 0)$, and the game Θ_y has infinitely many Nash equilibria, namely $\{(x_1, x_2) \in [-1, 1]^2 \mid x_1 = x_2\}$. Note that the function P satisfies assumption ($\mathcal{L}_1$) but $P(\cdot, y)$ is not differentiable on the set $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = x_2\}$.

Remark 2.4 The strong convexity of $P(\cdot, y)$ (instead of the additional uniformity requirement stated in ($\mathcal{L}_1$)) would suffice to prove Proposition 2.2, as well as the results in this and the next section. We chose to state at once the assumption ($\mathcal{L}_1$) for the same reasons discussed in Remark 2.1.

The existence of a unique solution to (BNP) is illustrated in the following result, where we denote by $NE(\Omega)$ the set of Nash equilibria of the game Ω.

Theorem 2.1. *Assume* $(\mathcal{F}_1)$*-*$(\mathcal{F}_2)$*,* $(\mathcal{G}_1)$*-*$(\mathcal{G}_2)$*,* $(\mathcal{L}_1)$ *and that the function* $P(\cdot, y)$ *is continuously differentiable on* X *for any* $y \in Y$*. Then the problem (BNP) has a unique solution* $(\bar{x}, \bar{y}) \in X \times Y$*, which satisfies:*

$$\bar{x} = r(\bar{y}), \quad \textit{where } \{\bar{y}\} = NE(\Omega).$$

Proof. By Proposition 2.1 the game Ω has a unique Nash equilibrium $\bar{y} \in Y$. Then the thesis follows by applying Proposition 2.2 for $y = \bar{y}$. $\square$

Remark 2.5 If assumption $(\mathcal{L}_1)$ and the continuous differentiability of $P(\cdot, y)$ in Theorem 2.1 are replaced with the following weaker assumption:

$(\mathcal{L}'_1)$ $P(\cdot, y)$ is convex and coercive on X for any $y \in Y$,

then the existence of solutions to (BNP) still holds. In fact, denoted by $\bar{y} \in Y$ the unique Nash equilibrium of Ω (see Proposition 2.1), assumption $(\mathcal{L}'_1)$ implies that $P(\cdot, \bar{y})$ has a minimizer over X (see, e.g., [5, Proposition 11.14]). Since any miminizer of $P(\cdot, \bar{y})$ is a Nash equilibrium of $\Theta_{\bar{y}}$ (by [35, Lemma 2.1]), the problem (BNP) has at least one solution. However, in general the uniqueness of a solution to (BNP) is no longer guaranteed. To show this, it is sufficient to consider the function P defined by

$$P(x, y) = \begin{cases} 0, & \text{if } x_1^2 + x_2^2 \leq 1 \\ x_1^2 + x_2^2 - 1, & \text{if } x_1^2 + x_2^2 > 1. \end{cases}$$

as the potential function in the upper-level game Θ_y. It satisfies assumption $(\mathcal{L}'_1)$, but $P(\cdot, y)$ has infinitely many minimizers over $\mathbb{R}^2$.

3 Preliminary results for the approximation

Here we show fundamental results on the games in the upper level and in the lower level which will be employed for the approximation of a solution to (BNP).

Let us start again with the lower-level game Ω, by displaying some properties of the functions b_j and b defined in (1) and (2), respectively.

Lemma 3.1 (Lemma 2.5 in [10])**.** *Assume* $(\mathcal{F}_1)$*-*$(\mathcal{F}_2)$ *and let* λ_j *and* λ *be the real numbers defined in* (3) *and* (5)*, respectively. Then*

(i) the function b_j *is continuously differentiable and Lipschitz continuous with Lipschitz constant no greater than* λ_j*, for any* $j \in J$*;*

(ii) the function b *is continuously differentiable and Lipschitz continuous with Lipschitz constant no greater than* λ*.*

The next results involves the affine combinations of b and of the identity map of Y_1, namely the function $t_\delta \colon Y_1 \to Y_1$ defined by

$$t_\delta(y_1) := \delta y_1 + (1 - \delta) b(y_1), \quad \text{where } \delta \in \mathbb{R}. \tag{8}$$

In particular, we illustrate contraction properties of t_δ when Ω is ratio-bounded. Such properties rely on results proved by the authors in [10, 11].

Lemma 3.2. *Assume* $(\mathcal{F}_1)$*-*$(\mathcal{F}_2)$ *and* $(\mathcal{G}_1)$*. Let* λ *be the real number defined in* (5) *and* (α,β) *such that* $(\mathcal{G}_1)$ *holds. Denote by* $\Delta^\lambda_{\alpha,\beta}$ *the interval*

$$\Delta^\lambda_{\alpha,\beta}=\begin{cases}\left]\frac{\lambda^2-1}{\lambda^2-2\alpha+1},1\right[, & \text{if } \lambda<1\\ \left]1,\frac{\lambda^2-1}{\lambda^2-2\alpha+1}\right[, & \text{if } \alpha>1\\ \left]\frac{\lambda^2-1}{\lambda^2-2\beta+1},1\right[, & \text{if } \lambda\geq 1 \text{ and } \beta<1.\end{cases}\tag{9}$$

Then, the function t_δ *is a contraction for any* $\delta\in\Delta^\lambda_{\alpha,\beta}$*.*

Proof. See [10, Theorem 2.10(i)] and the proofs of [11, theorems 4.1(iii) and 4.4(iii)]. □

Lemma 3.3. *Assume* $(\mathcal{F}_1)$*-*$(\mathcal{F}_2)$ *and* $(\mathcal{G}_1)$*. Let* λ *be the real number defined in* (5) *and* (α,β) *such that* $(\mathcal{G}_1)$ *holds. Then, the contraction constant of* t_δ *is minimal for* $\delta=\nu$*, where*

$$\nu=\begin{cases}\frac{\lambda^2-\alpha}{\lambda^2-2\alpha+1}, & \text{if } \lambda<\sqrt{\alpha} \text{ or } 1<\alpha\leq\lambda\\ \frac{\lambda^2-\beta}{\lambda^2-2\beta+1}, & \text{if } \beta<0 \text{ or } (0\leq\beta<1 \text{ and } \sqrt{\beta}<\lambda)\\ 0, & \text{if } \sqrt{\alpha}\leq\lambda<1 \text{ or } \lambda\leq\sqrt{\beta}<1,\end{cases}\tag{10}$$

and the corresponding contraction constant is

$$\kappa=\begin{cases}\left(\frac{\lambda^2-\alpha^2}{\lambda^2-2\alpha+1}\right)^{1/2}, & \text{if } \lambda<\sqrt{\alpha} \text{ or } 1<\alpha\leq\lambda\\ \left(\frac{\lambda^2-\beta^2}{\lambda^2-2\beta+1}\right)^{1/2}, & \text{if } \beta<0 \text{ or } (0\leq\beta<1 \text{ and } \sqrt{\beta}<\lambda)\\ \lambda, & \text{if } \sqrt{\alpha}\leq\lambda<1 \text{ or } \lambda\leq\sqrt{\beta}<1.\end{cases}\tag{11}$$

Proof. See [10, Theorem 2.10(ii)] and the proofs of [11, theorems 4.2 and 4.4(iv)]. □

As regards the parametric upper-level game Θ_y, we illustrate smoothness properties of the function r defined in (7). Before showing the result, let us introduce the following assumption on the function P defined in $(\mathcal{G}_2)$.

$(\mathcal{L}_2)$ P is twice continuously differentiable on $X\times Y$, the Hessian matrix $D^2_xP(x,y)\in\mathbb{R}^{p\times p}$ is invertible for any $(x,y)\in X\times Y$ and

$$\rho:=\sup_{(x,y)\in X\times Y}\|[D^2_xP(x,y)]^{-1}\cdot D_y(D_xP)(x,y)\|\in\mathbb{R};\tag{12}$$

Lemma 3.4. *Assume* $(\mathcal{G}_2)$*,* $(\mathcal{L}_1)$ *and* $(\mathcal{L}_2)$*. Then the function* r *is continuously differentiable and Lipschitz continuous with Lipschitz constant no greater then* ρ*.*

Proof. In light of Proposition 2.2, the function r satisfies $D_xP(r(y),y)=0$ for any $y\in Y$. Hence, by assumption $(\mathcal{L}_2)$, the Implicit Function Theorem ensures that the function r is continuously differentiable on Y and that $Dr(y)=[D^2_xP(r(y),y)]^{-1}\cdot D_y(D_xP)(r(y),y)$ for any $y\in Y$. Moreover, from the Mean Value Inequality we have

$$\|r(y')-r(y'')\|\leq\sup_{y\in Y}\|Dr(y)\|\|y'-y''\|\leq\rho\|y'-y''\|\quad\text{for any } y',y''\in Y.$$

Therefore, r is Lipschitz continuous with Lipschitz constant no greater than ρ. □

4 Numerical approximation

In order to numerically approximate a solution to (BNP), in this section we define the bilevel local variation method (BLVM for short) and we show its global convergence and related error estimations. Such a method exploits an optimization technique introduced in [15] for variational problems, called the local variation method (LVM for short). For the sake of completeness, we preliminarily illustrate the LVM and recall its associate convergence and errors estimation results.

4.1 The local variation method LVM

The LVM is a direct unconstrained optimization method that allows both to approach the unique minimizer of a strongly convex function and to obtain an estimation of the distance between the approximation calculated and the (exact) minimizer, by using only the values of the function.
The LVM is illustrated in Algorithm 1 for a function $g\colon \mathbb{R}^N \to \mathbb{R}$. The well-definedness of Algo-

---

Algorithm 1: Local Variation Method (LVM)

Data: Function $g\colon \mathbb{R}^N \to \mathbb{R}$, range $\epsilon > 0$, initial point $z_0^\epsilon \coloneqq (z_{0,1}^\epsilon, \dots, z_{0,N}^\epsilon) \in \mathbb{R}^N$.
Result: Stable point z^ϵ of g.

1 **begin**
2 $\quad k \leftarrow 0$;
3 $\quad$ **repeat**
4 $\quad\quad k \leftarrow k+1$;
5 $\quad\quad$ **for** $i = 1$ **to** N **do**
6 $\quad\quad\quad$ define:
7 $\quad\quad\quad \Delta_{k,i} \coloneqq g(z_{k,1}^\epsilon, \dots, z_{k,i-1}^\epsilon, z_{k-1,i}^\epsilon, z_{k-1,i+1}^\epsilon, \dots, z_{k-1,N}^\epsilon)$,
8 $\quad\quad\quad \Delta_{k,i}^+ \coloneqq g(z_{k,1}^\epsilon, \dots, z_{k,i-1}^\epsilon, z_{k-1,i}^\epsilon + \epsilon, z_{k-1,i+1}^\epsilon, \dots, z_{k-1,N}^\epsilon)$,
9 $\quad\quad\quad \Delta_{k,i}^- \coloneqq g(z_{k,1}^\epsilon, \dots, z_{k,i-1}^\epsilon, z_{k-1,i}^\epsilon - \epsilon, z_{k-1,i+1}^\epsilon, \dots, z_{k-1,N}^\epsilon)$;
10 $\quad\quad\quad$ find $\underset{\{z_{k-1,i}^\epsilon, z_{k-1,i}^\epsilon+\epsilon, z_{k-1,i}^\epsilon-\epsilon\}}{\operatorname{Arg\,min}} \{\Delta_{k,i}, \Delta_{k,i}^+, \Delta_{k,i}^-\}$ and denote it by $z_{k,i}^\epsilon$;
11 $\quad\quad$ **end**
12 $\quad\quad z_k^\epsilon \coloneqq (z_{k,1}^\epsilon, z_{k,2}^\epsilon, \dots, z_{k,N}^\epsilon)$;
13 $\quad$ **until** $g(z_k^\epsilon) \le g(z_{k,1}^\epsilon, \dots, z_{k,i-1}^\epsilon, z_{k,i}^\epsilon \pm \epsilon, z_{k,i+1}^\epsilon, \dots, z_{k,N}^\epsilon)$ *for any* $i \in \{1, \dots, N\}$;
14 $\quad z^\epsilon \coloneqq (z_{k,1}^\epsilon, z_{k,2}^\epsilon, \dots, z_{k,N}^\epsilon)$;
15 **end**

---

rithm 1 and the convergence towards a minimizer of g have been shown in [16]. For the sake of completeness we give below the proofs of such results.

Lemma 4.1 (Lemma 1.1 in [16])**.** *Assume that $g\colon \mathbb{R}^N \to \mathbb{R}$ is strongly convex and let $\epsilon > 0$. Then Algorithm 1 ends after a finite number of steps and gives a stable point $z^\epsilon \in \mathbb{R}^N$, that is a point satisfying*

$$g(z^\epsilon) \le g(z_1^\epsilon, \dots, z_{i-1}^\epsilon, z_i^\epsilon \pm \epsilon, z_{i+1}^\epsilon, \dots, z_N^\epsilon) \quad \textit{for any } i \in \{1, \dots, N\}. \tag{13}$$

Proof. By contradiction, suppose that Algorithm 1 does not end after a finite number of steps and let $(z_k^\epsilon)_k$ be the sequence where $z_k^\epsilon \coloneqq (z_{k,1}^\epsilon, z_{k,2}^\epsilon, \dots, z_{k,N}^\epsilon)$ is the vector obtained at line 12

of Algorithm 1. Then, the sequence $(z_k^\epsilon)_k$ is necessarily bounded. In fact, if it was not bounded, $\lim_{k\to+\infty}\|z_k^\epsilon\| = +\infty$ would imply $\lim_{k\to+\infty} g(z_k^\epsilon) = +\infty$, as g is strongly convex and so coercive. But this is not possible since $(g(z_k^\epsilon))_k$ is a decreasing sequence, by construction. Hence $(z_k^\epsilon)_k$ is bounded.
Let $C > 0$ such that $|z_{k,i}^\epsilon| \leq C$ for any $i \in \{1,\dots,N\}$ and $k \in \mathbb{N}$. Consequently, since $z_{k,i}^\epsilon = z_{0,i}^\epsilon + m_{k,i}\epsilon$ for some $m_{k,i} \in \mathbb{Z}$, there exists $\bar{k} \in \mathbb{N}$ such that $z_k^\epsilon = z_{\bar{k}}^\epsilon$ for any $k > \bar{k}$ and $z_{\bar{k}}^\epsilon$ necessarily satisfies (13). Therefore, Algorithm 1 ends after a finite number of steps and gives a stable point. □

Note that the LVM belongs to the class of multidimensional search methods without using derivatives and, in particular, it can be recognized as a pattern search method with discrete steps (originally introduced in [27]; see for example [6, Section 8.5] for further discussion).

Theorem 4.1 (Theorem 2.1 in [16])**.** *Assume that $g\colon \mathbb{R}^N \to \mathbb{R}$ is continuously differentiable and strongly convex. Let $(\epsilon_n)_n \subseteq]0,+\infty[$ be a sequence decreasing to zero and z^{ϵ_n} be the stable point obtained by Algorithm 1 with range ϵ_n. Then, the sequence $(z^{\epsilon_n})_n$ converges to the unique maximizer of g and $(g(z^{\epsilon_n}))_n$ converges to the minimum of g.*

Proof. For the sake of readability, we show the result for $N = 2$. Fixed $n \in \mathbb{N}$, let $z^{\epsilon_n} = (z_1^{\epsilon_n}, z_2^{\epsilon_n})$ be the stable point obtained by applying the LVM to g with range ϵ_n, which is well-defined by Lemma 4.1, and $z^{min} = (z_1^{min}, z_2^{min})$ be the unique minimizer of g over $\mathbb{R}^2$, whose existence is ensured by the strong convexity of g (see, for example, [5, Corollary 11.16]).
In light of inequality (13) and the Mean Value Theorem,

$$\begin{aligned}
0 \leq g(z_1^{\epsilon_n} + \epsilon_n, z_2^{\epsilon_n}) - g(z_1^{\epsilon_n}, z_2^{\epsilon_n}) = \epsilon_n Dg_{z_1}(z_1^{\epsilon_n} + \gamma_1\epsilon_n, z_2^{\epsilon_n})\\
0 \leq g(z_1^{\epsilon_n} - \epsilon_n, z_2^{\epsilon_n}) - g(z_1^{\epsilon_n}, z_2^{\epsilon_n}) = -\epsilon_n Dg_{z_1}(z_1^{\epsilon_n} - \gamma_1'\epsilon_n, z_2^{\epsilon_n})\\
0 \leq g(z_1^{\epsilon_n}, z_2^{\epsilon_n} + \epsilon_n) - g(z_1^{\epsilon_n}, z_2^{\epsilon_n}) = \epsilon_n Dg_{z_2}(z_1^{\epsilon_n}, z_2^{\epsilon_n} + \gamma_2\epsilon_n)\\
0 \leq g(z_1^{\epsilon_n}, z_2^{\epsilon_n} - \epsilon_n) - g(z_1^{\epsilon_n}, z_2^{\epsilon_n}) = -\epsilon_n Dg_{z_2}(z_1^{\epsilon_n}, z_2^{\epsilon_n} - \gamma_2'\epsilon_n),
\end{aligned}$$

where $\gamma_1, \gamma_1', \gamma_2, \gamma_2' \in]0,1[$ are depending on ϵ_n. Hence

$$\begin{aligned}
&Dg_{z_1}(z_1^{\epsilon_n} + \gamma_1\epsilon_n, z_2^{\epsilon_n}) \geq 0, \quad Dg_{z_1}(z_1^{\epsilon_n} - \gamma_1'\epsilon_n, z_2^{\epsilon_n}) \leq 0,\\
&Dg_{z_2}(z_1^{\epsilon_n}, z_2^{\epsilon_n} + \gamma_2\epsilon_n) \geq 0, \quad Dg_{z_2}(z_1^{\epsilon_n}, z_2^{\epsilon_n} - \gamma_2'\epsilon_n) \leq 0.
\end{aligned} \tag{14}$$

Now, note that the sequence $(g(z^{\epsilon_n}))_n$ is decreasing by construction, so the sequence $(z^{\epsilon_n})_n$ is bounded (by arguing as in the proof of Lemma 4.1). Let $(z_1^{\epsilon_{n_k}})_k \subseteq (z_1^{\epsilon_n})_n$ and $(z_2^{\epsilon_{n_k}})_k \subseteq (z_2^{\epsilon_n})_n$ two subsequences which converge to $\bar{z}_1 \in \mathbb{R}$ and $\bar{z}_2 \in \mathbb{R}$ as k goes to $+\infty$, respectively. Thus, replacing ϵ_n with ϵ_{n_k} in (14) and letting k to $+\infty$, the continuity of the partial derivatives of g implies that

$$Dg_{z_1}(\bar{z}_1, \bar{z}_2) = 0 \quad \text{and} \quad Dg_{z_2}(\bar{z}_1, \bar{z}_2) = 0,$$

i.e. $(\bar{z}_1, \bar{z}_2)$ is a critical point for g. In light of the strong convexity of g, such a critical point is unique and coincides with the minimizer of g over $\mathbb{R}^2$, that is $\bar{z}_1 = z_1^{min}$ and $\bar{z}_2 = z_2^{min}$. The boundedness of $(z^{\epsilon_n})_n$ and the uniqueness of the minimizer of g guarantees that the whole sequence $(z^{\epsilon_n})_n$ converges to z^{min}. Finally, the last part of the thesis follows by the continuity of g. □

Before showing the error estimations of the LVM, we preliminarily recall that, if $g\colon \mathbb{R}^N \to \mathbb{R}$ is differentiable and $z \in \mathbb{R}^N$, the Taylor's theorem guarantees

$$\exists\, \mathcal{I}_z \subseteq \mathbb{R}^N \text{ s.t. } \quad g(z+h) - g(z) = \langle \nabla g(z), h\rangle + w(z,h) \quad \forall h \in \mathcal{I}_z, \tag{15}$$

where $\mathcal{I}_z$ is a neighbourhood of 0 depending on z, $\nabla g(z) \in \mathbb{R}^N$ is the gradient of g at z, and the remainder $w(z,h)$ satisfies $\lim_{h\to 0} w(z,h)/\|h\|_{\mathbb{R}^N} = 0$. Moreover, if in addition g is strongly convex, then there exists $m > 0$ such that

$$g(z'') - g(z') \geq \langle \nabla g(z'), z'' - z'\rangle + m\|z'' - z'\|^2, \tag{16}$$

for any $x', x'' \in \mathbb{R}^N$. The error estimations of the LVM illustrated in the next result are obtained by exploiting the proofs of [16, Theorem 3.1] and [36, Theorem 2.3].

Proposition 4.1. *Assume that $g\colon \mathbb{R}^N \to \mathbb{R}$ is differentiable and strongly convex and that there exist $C_1 > 0$, $C_0 \geq 0$ and $\tau > 1$ such that*

$$|w(z,h)| \leq C_1\|h\|^\tau + C_0\|h\|^{\tau+1} \quad \textit{for any } z \in \mathbb{R}^N \textit{ and } h \in \mathcal{I}_z, \tag{17}$$

where w and $\mathcal{I}_z$ are defined in (15). Let $\epsilon > 0$ and let z^ϵ be the stable point obtained by Algorithm 1. Then

$$\|z^\epsilon - z^{min}\| \leq \frac{\sqrt{N}(C_1 + \epsilon C_0)}{m}\epsilon^{\tau-1}, \tag{18}$$

where z^{min} is the unique minimizer of g over $\mathbb{R}^N$ and m is the constant related to the strong convexity of g, defined in (16).

Proof. Firstly, note that z^ϵ is well-defined in light of Lemma 4.1. Let $\{e_1, \dots, e_N\}$ be the standard basis of $\mathbb{R}^N$ and let us fix $i \in \{1, \dots, N\}$. Since z^ϵ verifies (13), by (15) we have

$$\begin{aligned} 0 &\leq g(z^\epsilon - \epsilon e_i) - g(z^\epsilon) = -\epsilon\langle \nabla g(z^\epsilon), e_i\rangle + w(z^\epsilon, -\epsilon e_i), \\ 0 &\leq g(z^\epsilon + \epsilon e_i) - g(z^\epsilon) = \epsilon\langle \nabla g(z^\epsilon), e_i\rangle + w(z^\epsilon, \epsilon e_i). \end{aligned} \tag{19}$$

So, by (19) and (17)

$$\begin{aligned} \langle \nabla g(z^\epsilon), e_i\rangle &\leq \frac{w(z^\epsilon, -\epsilon e_i)}{\epsilon} \leq \frac{C_1\|-\epsilon e_i\|^\tau + C_0\|-\epsilon e_i\|^{\tau+1}}{\epsilon}, \\ \langle \nabla g(z^\epsilon), e_i)\rangle &\geq -\frac{w(z^\epsilon, \epsilon e_i)}{\epsilon} \geq -\frac{C_1\|\epsilon e_i\|^\tau + C_0\|\epsilon e_i\|^{\tau+1}}{\epsilon}. \end{aligned} \tag{20}$$

Since $\|e_i\| = 1$, we get

$$|\langle \nabla g(z^\epsilon), e_i)\rangle| \leq \epsilon^{\tau-1}(C_1 + \epsilon C_0). \tag{21}$$

Let m be the constant related to the strong convexity of g, as defined in (16). In light of the definition of z^{min} and inequalities (16) and (21), then

$$
\begin{aligned}
m\|z^{min}-z^{\epsilon}\|^2 &\leq [g(z^{min})-g(z^{\epsilon})]-\langle\nabla g(z^{\epsilon}), z^{min}-z^{\epsilon}\rangle \\
&\leq |\langle\nabla g(z^{\epsilon}), z^{min}-z^{\epsilon})\rangle| \\
&= \left|\langle\nabla g(z^{\epsilon}), \sum_{i=1}^{N}(z^{min}-z^{\epsilon})_i e_i)\rangle\right| \\
&\leq \sum_{i=1}^{N}|(z^{min}-z^{\epsilon})_i||\langle\nabla g(z^{\epsilon}), e_i\rangle| \\
&\leq \epsilon^{\tau-1}(C_1+\epsilon C_0)\|z^{min}-z^{\epsilon}\|_1 \\
&\leq \sqrt{N}(C_1+\epsilon C_0)\epsilon^{\tau-1}\|z^{min}-z^{\epsilon}\|,
\end{aligned}
$$

where $\|\cdot\|_1$ is the 1-norm of $\mathbb{R}^N$ and the last inequality follows from the equivalence of norms in $\mathbb{R}^N$, more precisely from the inequality $\|z\|_p \leq N^{(1/p-1/q)}\|z\|_q$ holding for any $z \in \mathbb{R}^N$ and $p, q \in [1, +\infty[$. Therefore, the error estimation in (18) is proved and the proof is complete. □

A further result on the error estimations of the LVM can be derived from Proposition 4.1 by means of the Taylor's theorem with Lagrange's form of the remainder.

Corollary 4.1. *Assume that $g\colon \mathbb{R}^N \to \mathbb{R}$ is twice differentiable and strongly convex and that there exists $C > 0$ such that*

$$\|D^2 g(z)\| \leq C \quad \text{for any } z \in \mathbb{R}^N. \tag{22}$$

Let $\epsilon > 0$ and let z^{ϵ} be the stable point obtained by Algorithm 1. Then

$$\|z^{\epsilon}-z^{min}\| \leq \frac{\sqrt{N}C}{2m}\epsilon,$$

where z^{min} is the unique minimizer of g over $\mathbb{R}^N$ and m is the constant related to the strong convexity of g, defined in (16).

4.2 The bilevel local variation method BLVM

In order to numerically approximate a solution to (BNP), now we define the bilevel local variation method (BLVM). At each step of BLVM the current strategy profile of all players is updated by employing the LVM presented in Section 4.1 three times: firstly, one time for each follower's payoff function f_j in the lower-level game Ω and then, going up, one time for the potential function P in the upper-level game Θ_y (where y comes from the approximation obtained in the lower-level game).

Assumed that (BNP) satisfies $(\mathcal{F}_1)$-$(\mathcal{F}_2)$, $(\mathcal{G}_1)$-$(\mathcal{G}_2)$ and $(\mathcal{L}_1)$, the BLVM is illustrated in Algorithm 2, where ν is the real number defined in (10). Figure 1 provides a graphical representation of Algorithm 2, by highlighting the order whereby the elements of the sequence are generated. In the next result, the global convergence of the bilevel local variation method towards a solution to (BNP) is shown.

Algorithm 2: Bilevel Local Variation Method (BLVM)

Data: Initial range $\epsilon_0 > 0$, initial point $(\tilde{w}_0, \tilde{u}_0, v_0) \in X \times Y_1 \times Y_2$.

Result: Numerical approximation of the solution to (BNP).

```
begin
    Apply Algorithm 1 to the function f_1(·, v_0) with range ε_0, initial point ũ_0 and get the
      stable point u_0^*;
    Apply Algorithm 1 to the function P(·, u_0^*, v_0) with range ε_0, initial point w̃_0 and get
      the stable point w_0^*;
    n ← 0;
    repeat
        n ← n + 1;
        Apply Algorithm 1 to the function f_2(u_{n−1}^*, ·) with range ε_0/2^n, initial point v_{n−1}^*
          and get the stable point v_n^*;
        Apply Algorithm 1 to the function f_1(·, v_n^*) with range ε_0/2^n, initial point u_{n−1}^* and
          get the stable point ũ_n^*;
        u_n^* := ν u_{n−1}^* + (1 − ν) ũ_n^*;
        Apply Algorithm 1 to the function P(·, u_n^*, v_n^*) with range ε_0/2^n, initial point w_{n−1}^*
          and get the stable point w_n^*;
end
```

Theorem 4.2. *Assume ($\mathcal{F}_1$)-($\mathcal{F}_2$), ($\mathcal{G}_1$)-($\mathcal{G}_2$), ($\mathcal{L}_1$)-($\mathcal{L}_2$), and that there exist $A > 0$, $B_1 > 0$ and $B_2 > 0$ such that*

$$\|D^2_{y_1} f_1(y)\| \leq B_1 \text{ for any } y \in Y, \tag{23a}$$

$$\|D^2_{y_2} f_2(y)\| \leq B_2 \text{ for any } y \in Y, \tag{23b}$$

$$\|D^2_{x} P(x, y)\| \leq A \text{ for any } (x, y) \in X \times Y. \tag{23c}$$

Let $\epsilon_0 > 0$ and $(\tilde{w}_0, \tilde{u}_0, v_0) \in X \times Y_1 \times Y_2$. Then, the sequence $(w_n^*, u_n^*, v_n^*)_n \subseteq X \times Y_1 \times Y_2$ generated by Algorithm 2 is convergent to the unique solution to the problem (BNP).

Proof. Preliminarily, note that the sequence $(w_n^*, u_n^*, v_n^*)_n \subseteq X \times Y_1 \times Y_2$ is well-defined and that problem (BNP) has a unique solution $(\bar{x}, \bar{y}_1, \bar{y}_2)$ by Theorem 2.1.
We start by proving the convergence of the sequence $(u_n^*)_n$ to $\bar{y}_1$. To show this, let us define the following auxiliary points

$$v_n := b_2(u_{n-1}) \in Y_2 \tag{24a}$$

$$u_n := \nu u_{n-1} + (1 - \nu) b_1(v_n) = t_\nu(u_{n-1}) \in Y_1 \tag{24b}$$

$$z_n := b_2(u_{n-1}^*) \in Y_2 \tag{24c}$$

$$\tilde{s}_n := b_1(z_n) \in Y_1 \tag{24d}$$

$$s_n := \nu u_{n-1}^* + (1 - \nu)\tilde{s}_n = t_\nu(u_{n-1}^*) \in Y_1 \tag{24e}$$

$$\tilde{h}_n := b_1(v_n^*) \in Y_1 \tag{24f}$$

$$h_n := \nu u_{n-1}^* + (1 - \nu)\tilde{h}_n \in Y_1, \tag{24g}$$

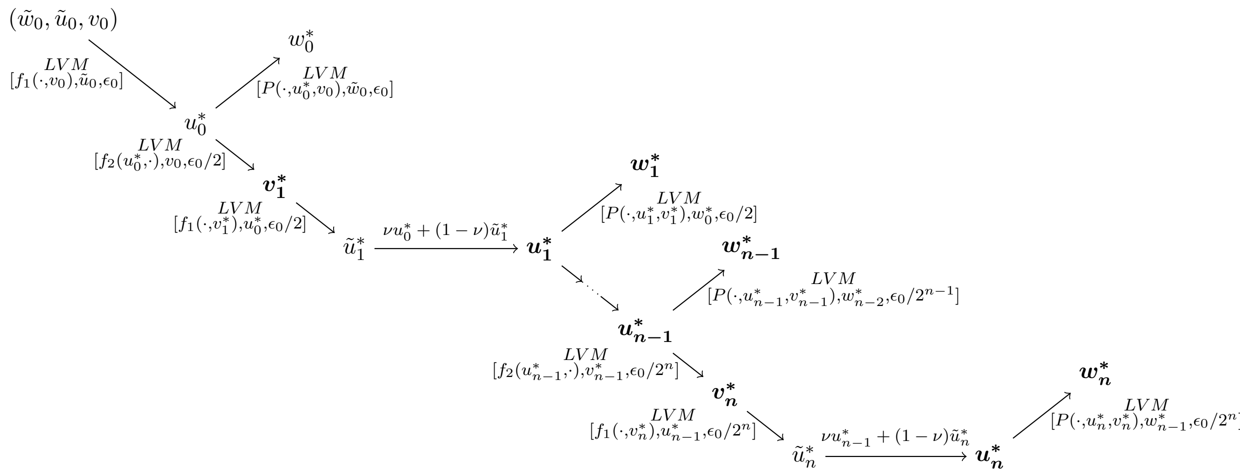

Figure 1: Scheme of Algorithm 2

for any $n \in \mathbb{N}$, where the functions b_1 and b_2 are defined in (1), $u_0 := b_1(v_0)$, u_0^* is defined in line 2 of Algorithm 2 and t_ν is defined in (8) (the connections among such auxiliary points and the sequences $(u_n^*)_n$ and $(v_n^*)_n$ are depicted in Figure 2). For any $n \in \mathbb{N}$

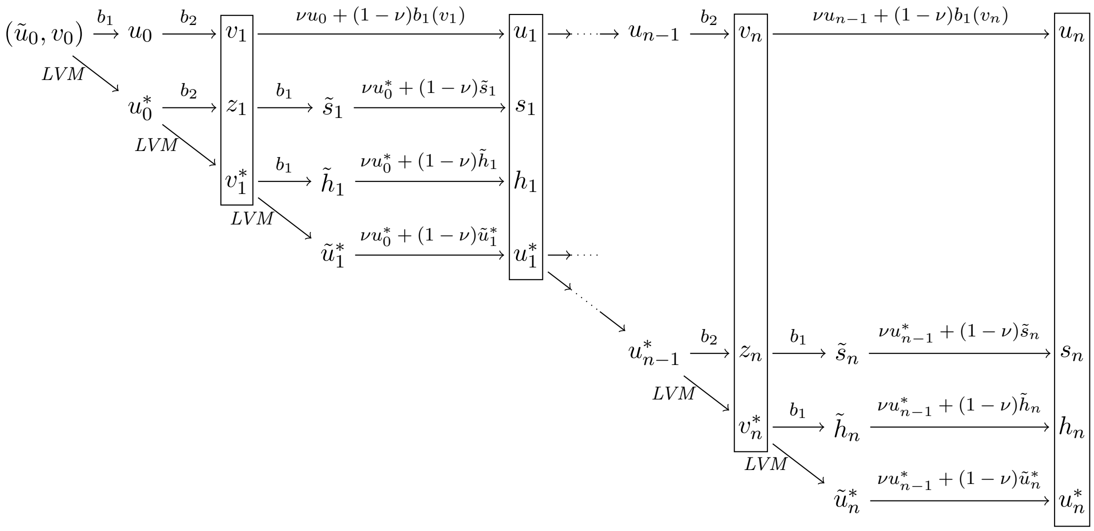

Figure 2: Representation of $v_k, u_k, z_k, \tilde{s}_k, s_k, \tilde{h}_k, h_k$, for $k = 1, \dots, n$.

$$\begin{aligned} \|u_n^* - \bar{y}_1\| &\leq \|u_n^* - u_n\| + \|u_n - \bar{y}_1\| \\ &\leq \|u_n^* - h_n\| + \|h_n - s_n\| + \|s_n - u_n\| + \|u_n - \bar{y}_1\|. \end{aligned} \tag{25}$$

Let us analyze the last four addends in the right-hand side of (25).

1. By definition of u_n^* in line 9 of Algorithm 2 and by (24g), we get

$$\|u_n^* - h_n\| = |1 - \nu| \|\tilde{u}_n^* - \tilde{h}_n\|. \tag{26}$$

Note that $\tilde{u}_n^*$ is the approximation of the minimizer of $f_1(\cdot, v_n^*)$ over Y_1 generated by applying Algorithm 1 to $f_1(\cdot, v_n^*)$ with initial point u_{n-1}^* and range $\epsilon_0/2^n$ (as represented in Figure 1), whereas

$\tilde{h}_n$ is actually such a minimizer, by (24f). So, in light of assumption (23a), from Corollary 4.1 we have

$$\|\tilde{u}_n^* - \tilde{h}_n\| \leq \frac{\sqrt{q_1}B_1}{2^{n+1}m_1}\epsilon_0, \tag{27}$$

where m_1 is the constant related to the concavity of f_1, as defined in assumption $(\mathcal{F}_1)$.

2. In light of (24d)–(24g) and Lemma 3.1(i), we have

$$\begin{aligned} \|h_n - s_n\| =& |1-\nu|\|\tilde{h}_n - \tilde{s}_n\| \\ =& |1-\nu|\|b_1(v_n^*) - b_1(z_n)\| \leq \lambda_1|1-\nu|\|v_n^* - z_n\|. \end{aligned} \tag{28}$$

Similarly to the previous case, v_n^* is the approximation of the minimizer of $f_2(u_{n-1}^*, \cdot)$ over Y_2 come up by applying Algorithm 1 to $f_2(u_{n-1}^*, \cdot)$ with initial point v_{n-1}^* and range $\epsilon_0/2^n$ (as represented in Figure 1), whereas z_n is effectively such a minimizer, by (24c). In light of assumption (23b), from Corollary 4.1 it follows that

$$\|v_n^* - z_n\| \leq \frac{\sqrt{q_2}B_2}{2^{n+1}m_2}\epsilon_0, \tag{29}$$

where m_2 is the constant related to the concavity of f_2, as defined in assumption $(\mathcal{F}_1)$.

3. By (24b) and (24e), we get

$$\|s_n - u_n\| = \|t_\nu(u_{n-1}^*) - t_\nu(u_{n-1})\|.$$

Since the function t_ν is a contraction and the corresponding contraction constant is κ defined in (11), from Lemma 3.3 we have

$$\|s_n - u_n\| \leq \kappa\|u_{n-1}^* - u_{n-1}\|. \tag{30}$$

Hence, from (26)-(30), we have

$$\|u_n^* - u_n\| \leq |1-\nu|\frac{\sqrt{q_1}B_1}{2^{n+1}m_1}\epsilon_0 + \lambda_1|1-\nu|\frac{\sqrt{q_2}B_2}{2^{n+1}m_2}\epsilon_0 + \kappa\|u_{n-1}^* - u_{n-1}\|. \tag{31}$$

Let $d_n = \|u_n^* - u_n\|$ for any $n \in \mathbb{N} \cup \{0\}$ and $D = |1-\nu|\left[\frac{\sqrt{q_1}B_1}{m_1} + \lambda_1\frac{\sqrt{q_2}B_2}{m_2}\right]\epsilon_0$. Then by (31) it follows that

$$d_n \leq \kappa d_{n-1} + \frac{D}{2^{n+1}} \leq \kappa\left[\kappa d_{n-2} + \frac{D}{2^n}\right] + \frac{D}{2^{n+1}} \leq \cdots \leq \kappa^n d_0 + \frac{D}{2}\sum_{\sigma=0}^{n-1}\frac{\kappa^\sigma}{2^{n-\sigma}}. \tag{32}$$

The summation $\sum_{\sigma=0}^{n}\frac{\kappa^\sigma}{2^{n-\sigma}}$ is the n-th term of the Cauchy product of the two series $\sum_{i=0}^{+\infty}\kappa^i$ and $\sum_{j=0}^{+\infty}\frac{1}{2^j}$, that is

$$\left(\sum_{i=0}^{+\infty}\kappa^i\right) \cdot_c \left(\sum_{j=0}^{+\infty}\frac{1}{2^j}\right) = \sum_{n=0}^{+\infty}\sum_{\sigma=0}^{n}\frac{\kappa^\sigma}{2^{n-\sigma}}, \tag{33}$$

where $\cdot_c$ denotes the Cauchy product. The two series in the left-hand side of (33) are geometric series with ratio less than 1 (recall that $\kappa < 1$ in light of Lemma 3.3), so they are convergent. Therefore, in light of the Cauchy theorem (see, for example, [23, Theorem 160]), the series in the right-hand side of (33) is convergent, hence $\lim_{n\to+\infty}\sum_{\sigma=0}^{n}\frac{\kappa^{\sigma}}{2^{n-\sigma}}=0$. Given the above and since $\lim_{n\to+\infty}\kappa^n=0$, by (32) we have

$$\lim_{n\to+\infty}\|u_n^*-u_n\|=\lim_{n\to+\infty}d_n=0. \tag{34}$$

4. In order to analyze the fourth addend, first recall that $\bar{y}_1$ is the unique fixed point of b in light of Proposition 2.1. Moreover, as $\nu\neq 1$, the set of fixed points of b coincides with the set of fixed points of t_ν. Hence, $\bar{y}_1$ is the unique fixed point of t_ν, that is $\bar{y}_1=t_\nu(\bar{y}_1)$. So, by (24b) and Lemma 3.3 we get

$$\|u_n-\bar{y}_1\|=\|t_\nu(u_{n-1})-t_\nu(\bar{y}_1)\|\leq\kappa\|u_{n-1}-\bar{y}_1\|\leq\cdots\leq\kappa^{n-1}\|u_1-\bar{y}_1\|,$$

and, consequently, as $\kappa<1$:

$$\lim_{n\to+\infty}\|u_n-\bar{y}_1\|=0. \tag{35}$$

Therefore, in light of (25),(34) and (35), the sequence $(u_n^*)_n$ converges to $\bar{y}_1$.

Now, we prove that the sequence $(v_n^*)_n$ converges to $\bar{y}_2$. For any $n\in\mathbb{N}$

$$\|v_n^*-\bar{y}_2\|\leq\|v_n^*-z_n\|+\|z_n-v_n\|+\|v_n-\bar{y}_2\|. \tag{36}$$

Let us analyze the three terms in the right-hand side of (36).

1. By (29), we get

$$\lim_{n\to+\infty}\|v_n^*-z_n\|\leq\lim_{n\to+\infty}\frac{\sqrt{q_2}B_2}{2^{n+1}m_2}\epsilon_0=0. \tag{37}$$

2. From (24a), (24c), Lemma 3.1(i) and (34) it follows that

$$\lim_{n\to+\infty}\|z_n-v_n\|=\lim_{n\to+\infty}\|b_2(u_{n-1}^*)-b_2(u_{n-1})\|\leq\lambda_2\lim_{n\to+\infty}\|u_{n-1}^*-u_{n-1}\|=0. \tag{38}$$

3. Since $\bar{y}_2=b_2(\bar{y}_1)$ (see Proposition 2.1), by (24a), Lemma 3.1(i) and (35) we have

$$\lim_{n\to+\infty}\|v_n-\bar{y}_2\|=\lim_{n\to+\infty}\|b_2(u_{n-1})-b_2(\bar{y}_1)\|\leq\lambda_2\lim_{n\to+\infty}\|u_{n-1}-\bar{y}_1\|=0. \tag{39}$$

Hence, in light of (36)-(39) the sequence $(v_n^*)_n$ is convergent to $\bar{y}_2$.

Finally, we show that the sequence $(w_n^*)_n$ converges to $\bar{x}$. For any $n\in\mathbb{N}$,

$$\|w_n^*-\bar{x}\|_X\leq\|w_n^*-r(u_n^*,v_n^*)\|+\|r(u_n^*,v_n^*)-r(u_n,v_n)\|+\|r(u_n,v_n)-\bar{x}\|, \tag{40}$$

where r is the function defined in (7). Let us analyze the three terms in the right-hand side of (40).

1. By line 10 in Algorithm 2, w_n^* is the approximation of the minimizer of $P(\cdot,u_n^*,v_n^*)$ over X generated by applying Algorithm 1 to $P(\cdot,u_n^*,v_n^*)$ with initial point w_{n-1}^* and range $\epsilon_0/2^n$ (as represented in Figure 1), whereas $r(u_n^*,v_n^*)$ is actually such a minimizer. So, in light of assumption (23c), from Corollary 4.1 we get

$$\|w_n^*-r(u_n^*,v_n^*)\|\leq\frac{\sqrt{p}A}{2^{n+1}m_{P}}\epsilon_0, \tag{41}$$

where m_{P} is the constant related to the concavity of P, as defined in assumption $(\mathcal{L}_1)$.

2. From Lemma 3.4 it follows that

$$\|r(u_n^*, v_n^*) - r(u_n, v_n)\| \leq \rho\|(u_n^*, v_n^*) - (u_n, v_n)\| \leq \rho\|u_n^* - u_n\| + \rho\|v_n^* - v_n\|, \tag{42}$$

where ρ is defined in (12).

3. Since $\bar{x} = r(\bar{y}_1, \bar{y}_2)$, in light of Lemma 3.4 we have

$$\|r(u_n, v_n) - \bar{x}\| = \|r(u_n, v_n) - r(\bar{y}_1, \bar{y}_2)\| \leq \rho\|(u_n, v_n) - (\bar{y}_1, \bar{y}_2)\| \leq \rho\|u_n - \bar{y}_1\| + \rho\|v_n - \bar{y}_2\|. \tag{43}$$

By (37)-(39) and (34)-(35), the terms in the right-hand sides of (41)-(43) converge to 0 as n goes to $+\infty$. Therefore, by (40), the sequence $(w_n^*)_n$ converges to $\bar{x}$. □

Remark 4.1 Assumptions (23a)-(23c) in Theorem 4.2 can be weakened. In fact, let $R_j(y_j, a_j, y_{-j})$ and $R_P(x, a, y)$ be the remainders of the Taylor expansion of $f_j(\cdot, y_{-j})$ and $P(\cdot, y)$ at y_j and x, respectively, which satisfy

$$\begin{aligned} f_j(y_j + a_j, y_{-j}) &= f_j(y_j, y_{-j}) + \langle \nabla_{y_j} f_j(y_j, y_{-j}), a_j \rangle + R_j(y_j, a_j, y_{-j}) \quad \forall a_j \in \mathcal{U}^j_{y_j, y_{-j}} \text{ and } \forall j \in J \\ P(x + a, y) &= P(x, y) + \langle \nabla_x P(x, y), a \rangle + R_P(x, a, y) \quad \forall a \in \mathcal{U}_{x,y}, \end{aligned}$$

where $\mathcal{U}^j_{y_j, y_{-j}} \subseteq Y_j$ and $\mathcal{U}_{x,y} \subseteq X$ are neighbourhoods of 0 in Y_j and X, respectively. By replacing (23a)-(23c) with the following conditions:

$$\begin{aligned} \|R_1(y_1, a_1, y_2)\| &\leq A_1'\|a_1\|^{\tau'} + A_0'\|a_1\|^{\tau'+1} \quad \text{for any } y_1 \in Y_1,\ a_1 \in \mathcal{U}^1_{y_1, y_2},\ y_2 \in Y_2 \\ \|R_2(y_2, a_2, y_2)\| &\leq A_1''\|a_2\|^{\tau''} + A_0''\|a_2\|^{\tau''+1} \quad \text{for any } y_2 \in Y_2,\ a_2 \in \mathcal{U}^2_{y_2, y_1},\ y_1 \in Y_1 \\ \|R_P(x, a, y)\| &\leq A_1\|a\|^{\tau} + A_0\|a\|^{\tau+1} \quad \text{for any } x \in X,\ a \in \mathcal{U}_{x,y},\ y \in Y, \end{aligned}$$

where $A_1' > 0$, $A_1'' > 0$, $A_1 > 0$, $A_0' \geq 0$, $A_0'' \geq 0$, $A_0 \geq 0$, $\tau' > 1$, $\tau'' > 1$, $\tau > 1$, and by using Proposition 4.1 instead of Corollary 4.1 in the proof of Theorem 4.2, the convergence of Algorithm 2 still holds. We preferred to assume the stronger conditions (23a)-(23c) for mere reasons of readability.

In the next result, the error estimation of the sequence $(w_n^*, u_n^*, v_n^*)_n$ generated by Algorithm 2 is shown.

Theorem 4.3. *Suppose that the assumptions of Theorem 4.2 hold. Then, there exist $F, G \in \mathbb{R}$ such that*

$$\|(w_n^*, u_n^*, v_n^*) - (\bar{x}, \bar{y}_1, \bar{y}_2)\| \leq F\kappa^n + \frac{G}{2^n} \quad \textit{for any } n \in \mathbb{N},$$

where κ is defined in (11).

Proof. Let $n \in \mathbb{N}$, then

$$\|(w_n^*, u_n^*, v_n^*) - (\bar{x}, \bar{y}_1, \bar{y}_2)\| \leq \|w_n^* - \bar{x}\| + \|u_n^* - \bar{y}_1\| + \|v_n^* - \bar{y}_2\|. \tag{44}$$

We start by proving the error estimation for the sequence $(u_n^*)_n$. Since

$$\|u_n^* - \bar{y}_1\| \leq \|u_n^* - u_n\| + \|u_n - \bar{y}_1\|, \tag{45}$$

where u_n is defined in (24b), let us analyze the two terms in the right-hand side of (45).

1. In light of (32) we know that

$$\|u_n^* - u_n\| \leq \frac{D}{2} \sum_{\sigma=0}^{n-1} \frac{\kappa^\sigma}{2^{n-\sigma}} + \kappa^n \|u_0^* - u_0\|, \tag{46}$$

where $D = |1-\nu| \left[\frac{\sqrt{q_1}B_1}{m_1} + \lambda_1 \frac{\sqrt{q_2}B_2}{m_2}\right] \epsilon_0$. By definition of u_0^* in line 2 of Algorithm 2 and u_0 (recall that $u_0 = b_1(v_0)$), Corollary 4.1 ensures that

$$\|u_0^* - u_0\| \leq \frac{\sqrt{q_1}B_1}{2m_1} \epsilon_0. \tag{47}$$

So, exploiting the sum of the first n terms of geometric series of ratio $2k$ in (46), we have

$$\|u_n^* - u_n\| \leq \frac{D}{2^{n+1}} \left[\frac{1-(2\kappa)^n}{1-2\kappa}\right] + \frac{\sqrt{q_1}B_1}{2m_1} \epsilon_0 \kappa^n. \tag{48}$$

2. By (24b) and Lemma 3.3, for any $s \in \mathbb{N}$ we get

$$\begin{aligned} \|u_{n+s} - u_n\| &\leq \sum_{j=1}^{s} \|u_{n+j} - u_{n+j-1}\| \\ &= \sum_{j=1}^{s} \|t_\nu(u_{n+j-1}) - t_\nu(u_{n+j-2})\| \\ &\leq \sum_{j=1}^{s} \kappa^{n+j-1} \|u_1 - u_0\| \\ &= \frac{\kappa^n(1-\kappa^s)}{1-\kappa} \|u_1 - u_0\|. \end{aligned} \tag{49}$$

As s goes to $+\infty$, by definition of u_1 and u_0 and Lemma 3.1, from (49) it follows that

$$\begin{aligned} \|u_n - \bar{y}_1\| &\leq \frac{\kappa^n}{1-\kappa} \|u_1 - u_0\| \\ &= \frac{\kappa^n}{1-\kappa} \|(1-\nu)(b_1(v_1) - b_1(v_0))\| \\ &\leq \frac{\lambda_1 |1-\nu| \kappa^n}{1-\kappa} \|v_1 - v_0\| \\ &\leq \frac{\lambda_1 |1-\nu| \kappa^n}{1-\kappa} [\|v_1 - z_1\| + \|z_1 - v_1^*\| + \|v_1^* - v_0\|]. \end{aligned} \tag{50}$$

In light of (24a), (24c), Lemma 3.1(i) and (47) we have $\|y_1 - z_1\| \leq \frac{\lambda_2 \sqrt{q_1} B_1}{2m_1} \epsilon_0$. Moreover, by definition of v_1^* in Algorithm 2 and of z_1 in (24c), Corollary 4.1 guarantees that $\|v_1^* - z_1\| \leq \frac{\sqrt{q_2}B_2}{4m_2} \epsilon_0$. Hence, from (50) it follows that

$$\|u_n - \bar{y}_1\| \leq \frac{\lambda_1 |1-\nu| \kappa^n}{1-\kappa} \left[\frac{\lambda_2 \sqrt{q_1} B_1}{2m_1} \epsilon_0 + \frac{\sqrt{q_2} B_2}{4m_2} \epsilon_0 + \|v_1^* - v_0\|\right]. \tag{51}$$

Finally, by using (48) and (51), from (45) we get

$$
\begin{aligned}
\|u_n^* - \bar{y}_1\| &\leq \frac{D}{2^{n+1}}\left[\frac{1-(2\kappa)^n}{1-2\kappa}\right] + \frac{\sqrt{q_1}B_1}{2m_1}\epsilon_0\kappa^n \\
&\quad + \frac{\lambda_1|1-\nu|\kappa^n}{1-\kappa}\left[\frac{\lambda_2\sqrt{q_1}B_1}{2m_1}\epsilon_0 + \frac{\sqrt{q_2}B_2}{4m_2}\epsilon_0 + \|v_1^* - v_0\|\right] \\
&= F'\kappa^n + \frac{G'}{2^n},
\end{aligned} \tag{52}
$$

where we set

$$
F' = \frac{D}{2(1-2\kappa)} + \frac{\sqrt{q_1}B_1}{2m_1}\epsilon_0 + \frac{\lambda_1|1-\nu|}{1-\kappa}\left[\frac{\lambda_2\sqrt{q_1}B_1}{2m_1}\epsilon_0 + \frac{\sqrt{q_2}B_2}{4m_2}\epsilon_0 + \|v_1^* - v_0\|\right],\ G' = \frac{D}{2(1-2\kappa)}. \tag{53}
$$

Now, we show the error estimation for the sequence $(v_n^*)_n$. Recalling that $\bar{y}_2 = b_2(\bar{y}_1)$, by (29), (24c), Lemma 3.1(i) and (52) we have

$$
\begin{aligned}
\|v_n^* - \bar{y}_2\| &\leq \|v_n^* - z_n\| + \|z_n - \bar{y}_2\| \\
&\leq \frac{\sqrt{q_2}B_2}{2^{n+1}m_2}\epsilon_0 + \lambda_2\|u_{n-1}^* - \bar{y}_1\| \\
&\leq F''\kappa^n + \frac{G''}{2^n}
\end{aligned} \tag{54}
$$

where

$$
F'' = \frac{\lambda_2 F'}{\kappa} \quad \text{and} \quad G'' = 2\lambda_2 G' + \frac{\sqrt{q_2}B_2}{2m_2}\epsilon_0, \tag{55}
$$

with F' and G' given in (53).

Finally, we prove the error estimation for the sequence $(w_n^*)_n$. Recalling that $\bar{x} = r(\bar{y}_1, \bar{y}_2)$, by (41), Lemma 3.4, (52) and (54) it follows that

$$
\begin{aligned}
\|w_n^* - \bar{x}\| &\leq \|w_n^* - r(u_n^*, v_n^*)\| + \|r(u_n^*, v_n^*) - r(\bar{y}_1, \bar{y}_2)\| \\
&\leq \frac{\sqrt{p}A}{2^{n+1}m_P}\epsilon_0 + \rho[\|u_n^* - \bar{y}_1\| + \|v_n^* - \bar{y}_2\|] \\
&\leq F'''\kappa^n + \frac{G'''}{2^n}
\end{aligned} \tag{56}
$$

where

$$
F''' = \rho F'\left(1 + \frac{\lambda_2}{\kappa}\right) \quad \text{and} \quad G''' = \rho\left[(1+2\lambda_2)G' + \frac{\sqrt{q_2}B_2}{2m_2}\epsilon_0\right] + \frac{\sqrt{p}A}{2^{n+1}m_P}\epsilon_0, \tag{57}
$$

with F' and G' given in (53).

Therefore, by using (52), (54) and (56) in (44) we obtain

$$
\|(w_n^*, u_n^*, v_n^*) - (\bar{x}, \bar{y}_1, \bar{y}_2)\| \leq F\kappa^n + \frac{G}{2^n}, \tag{58}
$$

with $F = F' + F'' + F'''$ and $G = G' + G'' + G'''$. □

The error estimation proved in Theorem 4.3 allows to derive the rate and the order of convergence of the sequence $(w_n^*, u_n^*, v_n^*)_n$.

Proposition 4.2. *Suppose that the assumptions of Theorem 4.2 hold and let* $T = \min\{\kappa^{-1}, 2\}$*. Then the sequence* $(w_n^*, u_n^*, v_n^*)_n$ *exhibits* $O(T^{-n})$*-rate of convergence and converges R-linearly to* $(\bar{x}, \bar{y}_1, \bar{y}_2)$*.*

Proof. First it is worth noting that $T > 1$ since $\kappa \in]0,1[$ by Lemma 3.3. Denoted by $\zeta_n := (|F| + |G|)T^{-n}$ for any $n \in \mathbb{N}$, from Theorem 4.3 we have

$$\|(w_n^*, u_n^*, v_n^*) - (\bar{x}, \bar{y}_1, \bar{y}_2)\| \leq \zeta_n \quad \text{for any } n \in \mathbb{N},$$

so $(w_n^*, u_n^*, v_n^*)_n$ has $O(T^{-n})$-rate of convergence. Moreover, since $T > 1$,

$$\lim_{n\to+\infty} \zeta_n = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{\zeta_{n+1}}{\zeta_n} = \frac{1}{T} \in]0,1[,$$

that is, the sequence $(\|(w_n^*, u_n^*, v_n^*) - (\bar{x}, \bar{y}_1, \bar{y}_2)\|)_n$ is dominated by a sequence converging linearly to 0. Therefore, $(w_n^*, u_n^*, v_n^*)_n$ converges R-linearly to $(\bar{x}, \bar{y}_1, \bar{y}_2)$ (see, e.g., [39, pp. 28–30]). $\square$

Remark 4.2 The error estimation and the rate of convergence proved in Theorem 4.3 and Proposition 4.2 are essentially affected by the sequence of ranges considered in Algorithm 2, namely $\epsilon_0/2^n$ for any $n \in \mathbb{N}$. We point out that, for any decreasing sequence of ranges $(\epsilon_n)_n$ such that the series $\sum_{n=0}^{+\infty} \epsilon_n$ is convergent, the convergence of Algorithm 2 still holds. Hence, improvements in the error estimations and in the rates of convergence could be achieved by choosing suitable sequences of ranges $(\epsilon_n)_n$.

Remark 4.3 It is worth noting that Algorithm 2 can be also exploited as a numerical approximation scheme for the class of ratio-bounded games introduced in [11]. In fact, if one considers only the lower-level game Ω and deletes line 3 and line 10 in Algorithm 2, assumptions $(\mathcal{F}_1)$, $(\mathcal{F}_2)$ and $(\mathcal{G}_1)$ guarantee the convergence of the sequence $(u_n^*, v_n^*)_n$ towards the unique Nash equilibrium $(\bar{y}_1, \bar{y}_2)$ of Ω, as well as error estimation and rate of convergence analogous to the ones proved in Theorem 4.3 and Proposition 4.2.

References

[1] B. Allaz and J.-L. Vila. "Cournot competition, forward markets and efficiency". In: *J. Econ. Theory* 59 (1993), pp. 1–16. DOI: https://doi.org/10.1006/jeth.1993.1001.

[2] E. Allevi, D. Aussel, and R. Riccardi. "On an equilibrium problem with complementarity constraints formulation of pay-as-clear electricity market with demand elasticity". In: *J. Global Optim.* 70 (2018), pp. 329–346. DOI: https://doi.org/10.1007/s10898-017-0595-9.

[3] T. Başar and G. Olsder. *Dynamic Noncooperative Game Theory*. Society for Industrial and Applied Mathematics, 1998.

[4] J. F. Bard. "Optimality conditions for the bilevel programming problem". In: *Nav. Res. Logis. Q.* 31 (1984), pp. 13–26. DOI: https://doi.org/10.1002/nav.3800310104.

[5] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York: Springer, 2011. ISBN: 978-1-4419-9466-0.

[6] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming: theory and algorithms*. Hoboken: John Wiley & Sons, 2006.

[7] R. Brânzei, L. Mallozzi, and S. Tijs. "Supermodular games and potential games". In: *J. Math. Econ.* 39 (2003), pp. 39–49.

[8] W. Candler and R. Townsley. "A linear two-level programming problem". In: *Comput. Oper. Res.* 9 (1982), pp. 59–76. DOI: https://doi.org/10.1016/0305-0548(82)90006-5.

[9] J. B. Cardell, C. C. Hitt, and W. W. Hogan. "Market power and strategic interaction in electricity networks". In: *Resour. Energy Econ.* 19 (1997), pp. 109–137. DOI: https://doi.org/10.1016/S0928-7655(97)00006-7.

[10] F. Caruso, M. C. Ceparano, and J. Morgan. "An inverse-adjusted best response algorithm for Nash equilibria". In: *SIAM J. Optim.* 30 (2020), pp. 1638–1663.

[11] F. Caruso, M. C. Ceparano, and J. Morgan. "Best response algorithms in ratio-bounded games: convergence of affine relaxations to Nash equilibria". CSEF Working Paper No. 593. 2020. URL: https://ideas.repec.org/p/sef/csefwp/593.html.

[12] F. Caruso, M. C. Ceparano, and J. Morgan. "Uniqueness of Nash equilibrium in continuous two-player weighted potential games". In: *J. Math. Anal. Appl.* 459 (2018), pp. 1208–1221.

[13] F. Caruso, M. B. Lignola, and J. Morgan. "Regularization and approximation methods in Stackelberg games and bilevel optimization". In: *Bilevel Optimization: Advances and Next Challenges*. Ed. by S. Dempe and A. Zemkoho. Cham: Springer International Publishing, 2020, pp. 77–138. DOI: 10.1007/978-3-030-52119-6_4.

[14] M. C. Ceparano and J. Morgan. "Equilibrium selection in multi-leader-follower games with vertical information". In: *TOP* 25 (2017), pp. 526–543. DOI: 10.1007/s11750-017-0444-5.

[15] F.L. Chernous'ko. "A local variation method for the numerical solution of variational problems". In: *USSR Comput. Math. Math. Phys.* 5 (1965), pp. 234–242. DOI: 10.1016/0041-5553(65)90136-9.

[16] Y. Cherruault. "Une méthode directe de minimisation et applications". In: *Rev. Fr. Inform. Rech. Oper.* 2 (1968), pp. 31–52. URL: http://www.numdam.org/item/M2AN_1968__2_2_31_0/.

[17] Y. Cherruault and P. Loridan. "Méthodes pour la recherche de points de selle". In: *J. Math. Anal. Appl.* 42 (1973), pp. 522–535. DOI: 10.1016/0022-247X(73)90160-1.

[18] M. R. Crisci and J. Morgan. "Implementation and numerical results of an approximation method for constrained saddle point problems". In: *Intern. J. Comput. Math.* 15 (1984), pp. 163–179. DOI: 10.1080/00207168408803408.

[19] V. DeMiguel and H. Xu. "A stochastic multiple-leader Stackelberg model: analysis, computation, and application". In: *Oper. Res.* 57 (2009), pp. 1220–1235. DOI: 10.1287/opre.1080.0686.

[20] S. Dempe. *Foundations of bilevel programming*. Springer US, 2002.

[21] S. Dempe and A. Zemkoho. *Bilevel Optimization*. Cham: Springer, 2020. DOI: 10.1007/978-3-030-52119-6.

[22] G. Facchini, F. van Megen, P. Borm, and S. Tijs. "Congestion models and weighted Bayesian potential games". In: *Theory Decis.* 42 (1997), pp. 193–206. DOI: 10.1023/A:1004991825894.

[23] G. H. Hardy. *Divergent Series*. Oxford: Oxford University Press, 1973. ISBN: 0198533098.

[24] R. Henrion, J. Outrata, and T. Surowiec. "Analysis of M-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market". In: *ESAIM Contr. Op. Ca. Va.* 18 (2012), pp. 295–317. DOI: 10.1051/cocv/2011003.

[25] M. Herty, S. Steffensen, and A. Thünen. "Solving quadratic multi-leader-follower games by smoothing the follower's best response". In: *Optim. Methods and Softw.* (2020). DOI: 10.1080/10556788.2020.1828412.

[26] B. F. Hobbs, C. B. Metzler, and J.-S. Pang. "Strategic gaming analysis for electric power systems: an MPEC approach". In: *IEEE T. Power Syst.* 15 (2000), pp. 638–645. DOI: 10.1109/59.867153.

[27] R. Hooke and T. A. Jeeves. ""Direct search" solution of numerical and statistical problems". In: *J. ACM* 8 (1961), 212–229. DOI: 10.1145/321062.321069.

[28] M. Hu and M. Fukushima. "Existence, uniqueness, and computation of robust Nash equilibria in a class of multi-leader-follower games". In: *SIAM J. Optim.* 23 (2013), pp. 894–916. DOI: 10.1137/120863873.

[29] X. Hu and D. Ralph. "Using EPECs to model bilevel games in restructured electricity markets with locational prices". In: *Oper. Res.* 55 (2007), pp. 809–827. DOI: 10.1287/opre.1070.0431.

[30] A. Koh and S. Shepherd. "Tolling, collusion and equilibrium problems with equilibrium constraints". In: *Eur. Transp.* (2010), pp. 3–22. URL: https://www.openstarts.units.it/bitstream/10077/6150/1/44NSKAD_KohShepherd.pdf.

[31] G. Leitmann. "On generalized Stackelberg strategies". In: *J. Optim. Theory Appl.* 26 (1978), pp. 637–643. DOI: 10.1007/BF00933155.

[32] P. Loridan and J. Morgan. "A theoretical approximation scheme for Stackelberg problems". In: *J. Optim. Theory Appl.* 61 (1989), pp. 95–110. DOI: 10.1007/BF00940846.

[33] P. Loridan and J. Morgan. "Approximation results for two-level optimization problem and application to penalty methods". Preprint n.52-1985. Dipartimento di Matematica e Applicazioni R. Caccioppoli, Università di Napoli. 1984. URL: https://doi.org/10.13140/RG.2.2.10836.55684.

[34] M. Maschler, E. Solan, and S. Zamir. "Game theory". In: *Cambridge: Cambridge University Press* (2013). DOI: doi:10.1017/CBO9780511794216.

[35] D. Monderer and L. S. Shapley. "Potential Games". In: *Games Econ. Behav.* 14 (1996), pp. 124 –143. DOI: https://doi.org/10.1006/game.1996.0044.

[36] J. Morgan. "Méthode directe de recherche du point de selle d'une fonctionnelle convexe-concave et application aux problèmes variationnels elliptiques avec deux contrôles antagonistes". (French). In: *Int. J. Comput. Math.* 4 (1974), pp. 143–175. DOI: 10.1080/00207167408803086.

[37] J. Morgan and P. Loridan. "Approximation of the Stackelberg Problem and Applications in Control Theory". In: *IFAC Proceedings Volumes* 18 (1986). 5th IFAC Workshop on Control Applications of Nonlinear Programming and Optimization (Capri, 1985), pp. 121–124. DOI: https://doi.org/10.1016/S1474-6670(17)60923-6.

[38] A. Neyman. "Correlated equilibrium and potential games". In: *Int. J. Game Theory* 26 (1997), pp. 223–227. DOI: 10.1007/BF01295851.

[39] J. Nocedal and S. Wright. *Numerical Optimization.* New York: Springer-Verlag, 2006. ISBN: 978-0-387-30303-1. DOI: 10.1007/978-0-387-40065-5.

[40] M. Pagnozzi and S. Piccolo. "Vertical separation with private contracts". In: *The Econ. J.* 122 (2011), pp. 173–207. DOI: 10.1111/j.1468-0297.2011.02471.x.

[41] D. Pal. "Cournot duopoly with two production periods and cost differentials". In: *J. Econ. Theory* 55 (1991), pp. 441–448. DOI: https://doi.org/10.1016/0022-0531(91)90050-E.

[42] J.-S. Pang and M. Fukushima. "Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games". In: *Comput. Manag. Sci.* 2 (2005), pp. 21–56. DOI: 10.1007/s10287-004-0010-0.

[43] J.B. Rosen. "Existence and uniqueness of equilibrium points for concave n-person games". In: *Econometrica* 33 (1965), pp. 520–534. DOI: 10.2307/1911749.

[44] G. Saloner. "Cournot duopoly with two production periods". In: *J. Econ. Theory* 42 (1987), pp. 183–187. DOI: https://doi.org/10.1016/0022-0531(87)90109-8.

[45] R. Selten. "Spieltheoretische behandlung eines oligopolmodells mit nachfrageträgheit: Teil i: Bestimmung des dynamischen preisgleichgewichts". In: *Z. Gesamte Staatswiss.* (1965), pp. 301–324. DOI: https://www.jstor.org/stable/40748884.

[46] H. D. Sherali. "A multiple leader Stackelberg model and analysis". In: *Oper. Res.* 32 (1984), pp. 390–404. DOI: 10.1287/opre.32.2.390.

[47] K. Shimizu and E. Aiyoshi. "A new computational method for Stackelberg and min-max problems by use of a penalty method". In: *IEEE T. Automat. Control* 26 (1981), pp. 460–466. DOI: 10.1109/TAC.1981.1102607.

[48] C.-L. Su. "Analysis on the forward market equilibrium model". In: *Oper. Res. Lett.* 35 (2007), pp. 74–82. DOI: https://doi.org/10.1016/j.orl.2006.01.006.

[49] C.-L. Su. "Equilibrium problems with equilibrium constraints: stationarities, algorithms, and applications". PhD thesis. Stanford, 2005. URL: https://web.stanford.edu/group/SOL/dissertations/clsu-thesis.pdf.

[50] M. Červinka. "Hierarchical Structures in Equilibrium Problems". PhD thesis. Prague, 2008. URL: https://dspace.cuni.cz/bitstream/handle/20.500.11956/19277/150001091.pdf?sequence=1&isAllowed=y.

[51] H. von Stackelberg. *Marktform und Gleichgewicht.* [Translation: D. Bazin, R. Hill and L. Urch. *Market Structure and Equilibrium.* Springer-Verlag Berlin Heidelberg, 2011]. Berlin: Verlag von Julius Springer, 1934.

[52] J. H. Wang, D. M. Chiu, and J. C. S. Lui. "A game–theoretic analysis of the implications of overlay network traffic on ISP peering". In: *Comput. Netw.* 52 (2008), pp. 2961–2974. DOI: https://doi.org/10.1016/j.comnet.2008.06.014.